Available online at www.sciencedirect.com

ScienceDirect

Journal of Approximation Theory 161 (2009) 65–87

Journal of Approximation Theory

www.elsevier.com/locate/jat

Orthogonal Laurent polynomials on the unit circle and snake-shaped matrix factorizations

Ruymán Cruz-Barroso[a], Steven Delvaux[b,*]

[a] *Department of Computer Science, K.U.Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium*
[b] *Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium*

Received 20 December 2007; received in revised form 25 April 2008; accepted 1 August 2008
Available online 5 November 2008

Communicated by Leonid Golinskii

Abstract

Let there be given a probability measure μ on the unit circle $\mathbb{T}$ of the complex plane and consider the inner product induced by μ. In this paper we consider the problem of orthogonalizing a sequence of monomials $\{z^{r_k}\}_k$, for a certain order of the $r_k \in \mathbb{Z}$, by means of the Gram–Schmidt orthogonalization process. This leads to a sequence of orthonormal Laurent polynomials $\{\psi_k\}_k$. We show that the matrix representation with respect to $\{\psi_k\}_k$ of the operator of multiplication by z is an infinite unitary or isometric matrix allowing a 'snake-shaped' matrix factorization. Here the 'snake shape' of the factorization is to be understood in terms of its graphical representation via sequences of little line segments, following an earlier work of S. Delvaux and M. Van Barel. We show that the shape of the snake is determined by the order in which the monomials $\{z^{r_k}\}_k$ are orthogonalized, while the 'segments' of the snake are canonically determined in terms of the Schur parameters for μ. Isometric Hessenberg matrices and unitary five-diagonal matrices (CMV matrices) follow as a special case of the presented formalism.

Keywords: Isometric Hessenberg matrix; Unitary five-diagonal matrix (CMV matrix); Givens transformation; Szegő polynomials; Orthogonal Laurent polynomials; Szegő quadrature formulas

* Corresponding author.
E-mail addresses: ruyman.cruzbarroso@cs.kuleuven.be (R. Cruz-Barroso), steven.delvaux@wis.kuleuven.be (S. Delvaux).

0021-9045/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jat.2008.08.004

1. Introduction

1.1. Isometric Hessenberg and unitary five-diagonal matrices

In recent years, there has been a lot of research activity on the topic of *unitary five-diagonal matrices*, also known as *CMV matrices*. These matrices have been used by researchers in various contexts, see e.g. [6–8,16,24,25,27,28,32].

Explicitly, the CMV matrix looks like

$$\mathcal{C} = \begin{bmatrix} \overline{\alpha_0} & \rho_0\overline{\alpha_1} & \rho_0\rho_1 & 0 & 0 & 0 & 0 & \dots \\ \rho_0 & -\alpha_0\overline{\alpha_1} & -\alpha_0\rho_1 & 0 & 0 & 0 & 0 & \dots \\ 0 & \rho_1\overline{\alpha_2} & -\alpha_1\overline{\alpha_2} & \rho_2\overline{\alpha_3} & \rho_2\rho_3 & 0 & 0 & \dots \\ 0 & \rho_1\rho_2 & -\alpha_1\rho_2 & -\alpha_2\overline{\alpha_3} & -\alpha_2\rho_3 & 0 & 0 & \dots \\ 0 & 0 & 0 & \rho_3\overline{\alpha_4} & -\alpha_3\overline{\alpha_4} & \rho_4\overline{\alpha_5} & \rho_4\rho_5 & \dots \\ 0 & 0 & 0 & \rho_3\rho_4 & -\alpha_3\rho_4 & -\alpha_4\overline{\alpha_5} & -\alpha_4\rho_5 & \dots \\ 0 & 0 & 0 & 0 & 0 & \rho_5\overline{\alpha_6} & -\alpha_5\overline{\alpha_6} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{1}$$

where α_k, $k = 0, 1, 2, \dots$ are complex numbers satisfying $|\alpha_k| < 1$ (the so-called *Schur parameters* or *Verblunsky coefficients*) and $\rho_k := \sqrt{1 - |\alpha_k|^2} \in (0, 1]$ are the so-called *complementary Schur parameters*. The matrix $\mathcal{C} = (c_{i,j})_{i,j\geq 0}$[1] in (1) can be seen to be unitary and five-diagonal, in the sense that $c_{i,j} = 0$ whenever $|i - j| > 2$. More precisely, the non-zero entries of $\mathcal{C}$ follow a kind of zigzag shape around the main diagonal.

The terminology 'CMV matrix' for the matrix in (1) originates from the book of Simon [27], who named these matrices after a 2003 paper by Cantero, Moral and Velázquez [7]. But this terminology is far from historically correct, since the latter paper [7] is in fact a rediscovery of facts which were already known by the numerical analysis community in the early 1990s; a survey of these early results can be found in the review paper by Watkins [33]; see also [28].

In the present paper, we prefer to avoid such historical discussions and we will therefore use the neutral term 'unitary five-diagonal matrix' to refer to these CMV matrices.

Unitary five-diagonal matrices have a number of interesting features, including the statement proven in the literature that (see further in this paper for more details) from all non-trivial classes of unitary matrices, unitary five-diagonals have the smallest bandwidth. Here the word 'non-trivial' refers to matrices which are not expressible as a direct sum of smaller matrices.

While this statement about the minimal bandwidth is certainly correct, it is a curious fact that this does *not* imply that unitary five-diagonal matrices are also *numerically* superior with respect to other non-trivial classes of unitary/isometric matrices. For example, another class of matrices which is often used in the literature is the class of *isometric Hessenberg matrices*, given explicitly by

[1] In the rest of the paper and for convenience with the notation, we will label the rows and columns of any matrix starting with index 0. As an example, the element $c_{1,1}$ in the matrix (1) will take the value $-\alpha_0\overline{\alpha_1}$.

$$\mathcal{H} = \begin{bmatrix} \overline{\alpha_0} & \rho_0\overline{\alpha_1} & \rho_0\rho_1\overline{\alpha_2} & \rho_0\rho_1\rho_2\overline{\alpha_3} & \rho_0\rho_1\rho_2\rho_3\overline{\alpha_4} & \dots \\ \rho_0 & -\alpha_0\overline{\alpha_1} & -\alpha_0\rho_1\overline{\alpha_2} & -\alpha_0\rho_1\rho_2\overline{\alpha_3} & -\alpha_0\rho_1\rho_2\rho_3\overline{\alpha_4} & \dots \\ 0 & \rho_1 & -\alpha_1\overline{\alpha_2} & -\alpha_1\rho_2\overline{\alpha_3} & -\alpha_1\rho_2\rho_3\overline{\alpha_4} & \dots \\ 0 & 0 & \rho_2 & -\alpha_2\overline{\alpha_3} & -\alpha_2\rho_3\overline{\alpha_4} & \dots \\ 0 & 0 & 0 & \rho_3 & -\alpha_3\overline{\alpha_4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{2}$$

Note that the matrix in (2) is of infinite dimension. This matrix is called *isometric* since its columns are orthonormal; a similar property for the rows is not guaranteed.

In (2) we use again the notation α_k, ρ_k to denote the Schur parameters and complementary Schur parameters, respectively. These are the same numbers as in matrix (1); see further. The matrix $\mathcal{H} = (h_{i,j})_{i,j\geq 0}$ in (2) is called *Hessenberg* since $h_{i,j} = 0$ whenever $i - j \geq 2$. Note however that the upper triangular part of this matrix is in general dense.

Now the point is that isometric Hessenberg matrices as in (2) are known to be *just as efficient* to manipulate as unitary five-diagonal matrices! Although this fact is known by numerical specialists, it seems that it is not so well-known in part of the theoretical community. Therefore, let us describe this now in somewhat more detail.

The naive idea would be that isometric Hessenberg matrices are 'inefficient' to work with since these matrices have a 'full' upper triangular part, in contrast to unitary five-diagonal matrices. But this would be a too quick conclusion. Having a better look at the problem, one can note that the upper triangular part of an isometric Hessenberg matrix is *rank structured* in the sense that each submatrix that can be taken out of the upper triangular part of such a matrix, has rank at most equal to 1. This can be easily verified using e.g. the explicit expressions of the entries of the matrix $\mathcal{H}$ in (2).

Going one step further, one can note that the rank structure in the upper triangular part of $\mathcal{H}$ is in fact a consequence of an even more structural theorem. Denote with $G_{k,k+1}$ a *Givens transformation* (also called Jacobi transformation)

$$G_{k,k+1} = \begin{bmatrix} I_k & 0 & 0 \\ 0 & \tilde{G}_{k,k+1} & 0 \\ 0 & 0 & I \end{bmatrix}, \tag{3}$$

where I_k and I denote identity matrices of sizes k and ∞, respectively, and where $\tilde{G}_{k,k+1}$ is a 2×2 unitary matrix positioned in rows and columns $\{k, k+1\}$. Thus the matrix $G_{k,k+1}$ differs from the identity matrix only by its entries in rows and columns $\{k, k+1\}$. Givens transformations can be considered as the most elementary type of unitary matrices. They can be used as building blocks to construct more general unitary matrices. Of interest for the present discussion is the fact that any (infinite) isometric Hessenberg matrix $\mathcal{H}$ allows a factorization as a product of Givens transformations in the form

$$\mathcal{H} = G_{0,1}G_{1,2}G_{2,3}G_{3,4}\dots. \tag{4}$$

This factorization must be understood in the sense that the principal $n \times n$ submatrices of $\mathcal{H}$ and $G_{0,1}G_{1,2}\dots G_{n-1,n}$ coincide for each n. This can be shown using only some basic linear algebra [15,17].

Applying this factorization to the matrix $\mathcal{H}$ in (2), one can actually specify this result by noting that the kth Givens transformation $G_{k,k+1}$ in (4) must have non-trivial part given by

$$\tilde{G}_{k,k+1} = \begin{bmatrix} \overline{\alpha_k} & \rho_k \\ \rho_k & -\alpha_k \end{bmatrix}. \tag{5}$$

In other words, the 'cosines' and 'sines' of the Givens transformations in (4) are nothing but the Schur parameters and complementary Schur parameters, respectively. This result was first established in the present context by Ammar, Gragg and Reichel [1].

Incidently, note that the Givens transformations in (5) are of a special form in the sense that they have real positive off-diagonal elements and determinant -1.

We also note the following finite-dimensional equivalent of (4): any unitary Hessenberg matrix $\mathcal{H}$ of size $n \times n$ allows a factorization in the form

$$\mathcal{H} = G_{0,1}G_{1,2}G_{2,3}G_{3,4}\dots G_{n-2,n-1}, \tag{6}$$

for suitable Givens transformations $G_{k,k+1}$, $k = 0, 1, \dots, n-2$.

The main point of (6) is that it shows that unitary Hessenberg matrices of size n can be compactly represented using only $O(n)$ parameters, just as is the case for unitary five-diagonal ones. Working with such an $O(n)$ matrix representation, the eigenvalue problem for unitary Hessenberg matrices can be solved numerically in a fast and accurate way; see the end of Section 4 for some references to eigenvalue computation algorithms in the literature. These algorithms can be canonically expressed in terms of the matrix factorization (6), i.e., in terms of the Schur parameters of the problem.

1.2. Graphical representation

In [13], a graphical notation was introduced where matrix factorizations with Givens transformations are represented via sequences of little line segments.

The graphical representation is obtained as follows. Let A be some arbitrary matrix (which will play no role in what follows) and suppose that we update $A \mapsto G_{k,k+1}A$. This means that the kth and $(k+1)$th row of A are replaced by linear combinations thereof, while the other rows of A are left unaltered. We can visualize this operation by drawing a vertical line segment on the left of the two modified rows of A.

One can then apply this idea in an iterative way. For example, when updating A by means of an operation $A \mapsto G_{k+1,k+2}G_{k,k+1}A$, one places first a vertical line segment on the left of rows $k, k+1$ (this deals with the update $A \mapsto G_{k,k+1}A$), and subsequently places a second vertical line segment on the left of the former one, this time at the height of rows $k+1, k+2$. We obtain in this way two successive vertical line segments. Clearly, any number of Givens transformations can be represented in such a way.

Now the key point is that we *identify* each $G_{k,k+1}$ with its corresponding vertical line segment. We hereby make abstraction of the matrix A on whose rows these operations were assumed to act. For example, the graphical representation of the factorization (6) with $n = 8$ is shown in Fig. 1.

Concerning Fig. 1, note that the top leftmost line segment in this figure (which is assumed to be placed at 'height' 0 and 1; cf. the indices on the left of the figure) corresponds to the leftmost factor $G_{0,1}$ in (6). Similarly, the second line segment corresponds to the factor $G_{1,2}$ in (6), and so on. We again emphasize that the line segments in Fig. 1 should be imagined as 'acting' on the rows of some (invisible) matrix A. See [13,14] for more applications of this graphical notation.

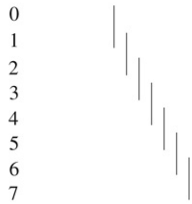

Fig. 1. The figure shows in a graphical way the decomposition as a product of Givens transformations of the unitary Hessenberg matrix $\mathcal{H}$ in (6) with $n = 8$.

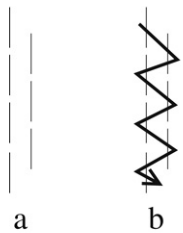

Fig. 2. The figure shows in a graphical way (a) the decomposition as a product of Givens transformations of the unitary five-diagonal matrix (7), (b) the 'snake shape' underlying this decomposition.

It is known that also unitary five-diagonal matrices allow a factorization as a product of Givens transformations. More precisely [1,4,7,33], the matrix $\mathcal{C}$ in (1) allows the factorization

$$\mathcal{C} = \begin{bmatrix} \overline{\alpha_0} & \rho_0 & 0 & 0 & 0 & \dots \\ \rho_0 & -\alpha_0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \overline{\alpha_2} & \rho_2 & 0 & \dots \\ 0 & 0 & \rho_2 & -\alpha_2 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & \overline{\alpha_1} & \rho_1 & 0 & 0 & \dots \\ 0 & \rho_1 & -\alpha_1 & 0 & 0 & \dots \\ 0 & 0 & 0 & \overline{\alpha_3} & \rho_3 & \dots \\ 0 & 0 & 0 & \rho_3 & -\alpha_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

which can be rewritten as

$$\mathcal{C} = (\dots G_{6,7}G_{4,5}G_{2,3}G_{0,1}) \cdot (G_{1,2}G_{3,4}G_{5,6}\dots), \tag{7}$$

where the $G_{k,k+1}$ are again defined by (3) and (5). Again, this factorization must be understood in the sense that the principal $n \times n$ submatrices of $\mathcal{C}$ and $G_{n-2,n-1}\dots G_{0,1} \cdot G_{1,2}\dots G_{n-1,n}$ for n even or $G_{n-1,n}\dots G_{0,1} \cdot G_{1,2}\dots G_{n-2,n-1}$ for n odd coincide for each n. The factorization (7) is represented graphically for $n = 8$ in Fig. 2.

Let us comment on Fig. 2. The leftmost series of line segments in Fig. 2(a) corresponds to the leftmost factor in (7). The order in which these Givens transformations are multiplied is clearly irrelevant; therefore we are allowed to place them all graphically aligned along the same vertical line. Similarly, the rightmost series of line segments in Fig. 2(a) corresponds to the rightmost factor in (7). To explain Fig. 2(b), imagine that one moves from the top to the bottom of the graphical representation. Then one can imagine a certain zigzag 'snake shape' underlying the factorization, which is shown in Fig. 2(b).

Note that in the above discussions we did not describe the way how isometric Hessenberg and unitary five-diagonal matrices arise in practice as matrix representations of a certain operator. At present, it will suffice to know that they are matrix representations of the *operator of multiplication by* z, acting on a function space generated by a certain sequence of orthonormal Laurent polynomials $\{\psi_k(z)\}_k$. This orthonormal sequence is obtained by applying the Gram–Schmidt orthogonalization process to the sequence of monomials

$$1, z, z^2, z^3, \ldots, \quad \text{and} \quad 1, z, z^{-1}, z^2, z^{-2}, \ldots, \tag{8}$$

for the isometric Hessenberg and unitary five-diagonal case, respectively.

1.3. Snake-shaped matrix factorizations

The aim of this paper it to carry the above observations one step further. We will show that with respect to a *general* sequence of orthonormal Laurent polynomials $\{\psi_k(z)\}_k$, obtained by orthogonalizing a *general* sequence of monomials (satisfying some conditions to be described in detail in Section 2.1), the operator of multiplication by z is represented by an infinite unitary or isometric matrix[2] allowing a snake-shaped matrix factorization. We will use the latter term to denote an infinite matrix product $\mathcal{S} = \prod_{k=0}^{\infty} G_{k,k+1}$, where the factors under the $\prod$-symbol are multiplied in a certain order. Here the ‘segments’ $G_{k,k+1}$ of the snake are canonically fixed in terms of the Schur parameters by means of (3) and (5), while the ‘shape’ of the snake, i.e., the order in which the $G_{k,k+1}$ are multiplied, will be determined by the order in which the monomials have been orthogonalized.

To fix the ideas, consider the sequence of monomials

$$1, z^{-1}, z, z^{-2}, z^2, z^3, z^{-3}, z^{-4}, z^4, z^5, \ldots. \tag{9}$$

With respect to the resulting sequence of orthonormal Laurent polynomials $\{\psi_k(z)\}_k$ (see Section 2.1 for details), the operator of multiplication by z will be described by a snake-shaped matrix factorization $\mathcal{S} = \mathcal{S}^{(\infty)}$. We claim that this factorization is built by means of the following recipe:

1. Considering the monomial $1 = z^0$ in the position 0 of (9), we initialize $\mathcal{S}^{(0)} := G_{0,1}$. Then we apply the following procedure for $k \geq 1$:
2. If the kth monomial in (9) has a *positive* exponent, we multiply the matrix with a new Givens transformation on the *right* by setting $\mathcal{S}^{(k)} := \mathcal{S}^{(k-1)} G_{k,k+1}$;
3. If the kth monomial in (9) has a *negative* exponent, we multiply the matrix with a new Givens transformation on the *left* by setting $\mathcal{S}^{(k)} := G_{k,k+1} \mathcal{S}^{(k-1)}$.

For the sequence of monomials (9), this recipe gives rise to the following series of iterate matrices $\mathcal{S}^{(k)}$:

$$\begin{aligned}
&\mathcal{S}^{(0)} = G_{0,1}, && \mathcal{S}^{(1)} = G_{1,2} \cdot G_{0,1}, \\
&\mathcal{S}^{(2)} = G_{1,2} \cdot G_{0,1} G_{2,3}, && \mathcal{S}^{(3)} = G_{3,4} G_{1,2} \cdot G_{0,1} G_{2,3} \\
&\mathcal{S}^{(4)} = G_{3,4} G_{1,2} \cdot G_{0,1} G_{2,3} G_{4,5}, && \mathcal{S}^{(5)} = G_{3,4} G_{1,2} \cdot G_{0,1} G_{2,3} G_{4,5} G_{5,6}, \\
&\mathcal{S}^{(6)} = G_{6,7} G_{3,4} G_{1,2} \cdot G_{0,1} G_{2,3} G_{4,5} G_{5,6}, && \ldots.
\end{aligned}$$

This leads to the final matrix factorization

[2] Note of caution: we will also consider certain cases where the subspace generated by the $\{\psi_k(z)\}_k$ is *not* invariant under the action of the operator of multiplication by z. In such cases, the above statement has to be formulated more carefully in order to make sure what the meaning is of the matrix $\mathcal{S}$; actually this matrix need not be unitary nor isometric then. For a precise statement we refer to the three cases distinguished at the beginning of Section 2.2, especially case 3.

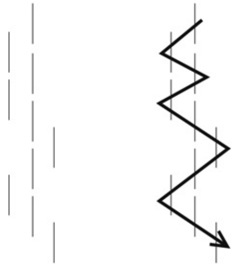

Fig. 3. The figure shows in a graphical way: (a) the decomposition as a product of Givens transformations of the matrix $\mathcal{S}$ in (10), (b) the 'snake shape' underlying this decomposition.

$$\mathcal{S} = \mathcal{S}^{(\infty)} = (\ldots G_{7,8}G_{6,7}G_{3,4}G_{1,2}) \cdot (G_{0,1}G_{2,3}G_{4,5}G_{5,6}G_{8,9}G_{9,10}\ldots). \tag{10}$$

This factorization is shown graphically in Fig. 3.

Let us comment on Fig. 3. The 'snake' in this figure was built by means of the following recipe:

1. Starting with a snake consisting of a single line segment $G_{0,1}$, we apply the following procedure for $k \geq 1$:
2. If the kth monomial in (9) has a *positive* exponent, the snake moves towards the *bottom right*, i.e., we add a new line segment on the bottom right of the snake;
3. If the kth monomial in (9) has a *negative* exponent, the snake moves towards the *bottom left*, i.e., we add a new line segment on the bottom left of the snake.

Of course, this recipe is nothing but a direct translation of the recipe that led us to the matrix factorization (10).

The reader should check that the above procedures are also valid for the isometric Hessenberg and for the unitary five-diagonal case (cf. (8) and Figs. 1 and 2).

1.4. Outline and contributions of the paper

The fact that the recipe in Section 1.3 leads to the correct matrix representation of the operator of multiplication by z with respect to the sequence of orthonormal Laurent polynomials $\{\psi_k(z)\}_k$ will be shown in Section 2. Our proof makes use of essentially three facts: (i) an observation of Cruz-Barroso et al. [12] (see also Watkins [33]) expressing the intimate connection between orthonormal Laurent polynomials and Szegő polynomials; (ii) the well-known Szegö recursion [30]; and (iii) an argument of Simon [28] using 'intermediary bases' in the isometric Hessenberg case. The full proof is however rather technical and requires some administrational book-keeping.

By factoring out a snake-shaped matrix product like (10), one can obtain explicit expressions for the entries of the matrix, generalizing the expansions in (1) and (2). This will be the topic of Section 3, where we will describe a graphical rule for determining the zero pattern of the matrix $\mathcal{S}$ as well as the shape of its non-zero elements.

Finally, in Section 4 we will briefly consider some connections between snake-shaped matrix factorizations and Szegő quadrature formulas. We will show that the known results involving isometric Hessenberg and unitary five-diagonal matrices can all be formulated in terms of a general snake-shaped matrix factorization $\mathcal{S}$, extending an observation of Ammar, Gragg and Reichel [1].

The remainder of this paper is organized as follows. Section 2 discusses some preliminaries about sequences of orthogonal Laurent polynomials on the unit circle and proves the main result about snake-shaped matrix factorizations. Section 3 discusses the entry-wise expansion of snake-shaped matrix factorizations. Finally, Section 4 considers the connection with the Szegő quadrature formulas.

To end this introduction, let us discuss the main contributions of this paper. It follows from the results presented here that isometric Hessenberg and unitary five-diagonal matrices can be considered as two extreme cases of a single mechanism, cf. the discussion in Section 1.3. In this way we obtain a unifying approach to some earlier results and estimates in the literature, see e.g. [6,8,12,27]. In addition, in the paper we provide graphical illustrations of the obtained matrix factorizations. These graphics lead to additional insight, explaining e.g. the term 'snake-shaped matrix factorization'. We feel that this might be an important conceptual contribution in its own respect.

2. Snake-shaped matrix factorizations: Main result

This section is devoted to the proof of our main result about snake-shaped matrix factorizations, showing how these occur as the matrix representation of the operator of multiplication by z with respect to a sequence of orthonormal Laurent polynomials. We start with some preliminaries.

2.1. Sequences of orthogonal Laurent polynomials on the unit circle

In this first subsection we fix some notations and conventions concerning orthogonal Laurent polynomials on the unit circle (see [6,10]-[12]). We denote by $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ the unit circle in the complex plane and by $\Lambda := \mathbb{C}[z, z^{-1}]$ the complex vector space of Laurent polynomials in the variable z. For a given order $n \in \mathbb{N}$ and an ordinary polynomial $p(z) = \sum_{k=0}^{n} c_k z^k$, we define its dual as $p^*(z) := z^n \overline{p(1/\overline{z})}$, or explicitly $p^*(z) = \sum_{k=0}^{n} \overline{c_{n-k}} z^k$. Here the bar denotes complex conjugation.

Throughout the paper, we shall be dealing with a finite positive non-discrete Borel measure μ supported on the unit circle $\mathbb{T}$ (which induces a measure on the interval $[-\pi, \pi]$ that we also denote by μ), normalized by the condition $\int_{-\pi}^{\pi} \mathrm{d}\mu(\theta) = 1$, i.e., a probability measure. As usual, the inner product induced by μ is given by

$$\langle f, g\rangle = \int_{-\pi}^{\pi} \overline{f\left(\mathrm{e}^{\mathrm{i}\theta}\right)} g(\mathrm{e}^{\mathrm{i}\theta}) \mathrm{d}\mu(\theta), \tag{11}$$

and the space of quadratically integrable functions with respect to the inner product (11) is denoted as $L_2^{\mu}(\mathbb{T})$.

For our purposes, we start constructing a sequence of subspaces of Laurent polynomials $\{\mathcal{L}_n\}_{n=0}^{\infty}$ satisfying

$$\mathcal{L}_0 := \mathrm{span}\{1\}, \qquad \dim(\mathcal{L}_n) = n + 1, \qquad \mathcal{L}_n \subset \mathcal{L}_{n+1}, \quad n \geq 1.$$

This can be done, by taking a sequence $\{p_n\}_{n=0}^{\infty}$ of non-negative integers such that $p_0 = 0$, $0 \leq p_n \leq n$ and $s_n := p_n - p_{n-1} \in \{0, 1\}$ for all $n \geq 1$. In what follows, a sequence $\{p_n\}_{n=0}^{\infty}$

satisfying these requirements will be called a *generating sequence*. Observe that in this case both $\{p_n\}_{n=0}^{\infty}$ and $\{n - p_n\}_{n=0}^{\infty}$ are non-negative non-decreasing sequences. Then, set

$$\mathcal{L}_n := \operatorname{span}\left\{z^j \ : \ -p_n \leq j \leq n - p_n\right\}$$

and set $\mathcal{L}_{-1} := \{0\}$ to be the trivial subspace. Observe that $\Lambda = \bigcup_{n=0}^{\infty} \mathcal{L}_n$ if and only if $\lim_{n\to\infty} p_n = \lim_{n\to\infty}(n - p_n) = \infty$ and that for all $n \geq 1$,

$$\mathcal{L}_n = \begin{cases} \mathcal{L}_{n-1} \oplus \operatorname{span}\{z^{n-p_n}\} & \text{if } s_n = 0, \\ \mathcal{L}_{n-1} \oplus \operatorname{span}\{z^{-p_n}\} & \text{if } s_n = 1. \end{cases}$$

Denote

$$\mathcal{L} := \overline{\bigcup_{n=0}^{\infty} \mathcal{L}_n}, \tag{12}$$

where $\overline{A}$ denotes the closure of A with respect to the norm induced by the inner product in $L_2^{\mu}(\mathbb{T})$. From the fact that the Laurent polynomials form a dense subset in $L_2^{\mu}(\mathbb{T})$, we have that $\mathcal{L} = L_2^{\mu}(\mathbb{T})$ if $\lim_{n\to\infty} p_n = \lim_{n\to\infty}(n - p_n) = \infty$. If this condition is violated, then $\mathcal{L}$ may be a *strict* subspace of $L_2^{\mu}(\mathbb{T})$.

By applying the Gram–Schmidt orthogonalization procedure to $\mathcal{L}_n$, an orthonormal basis $\{\psi_0(z), \ldots, \psi_n(z)\}$ can be obtained. If we repeat the process for each $n \geq 0$, a sequence $\{\psi_n(z)\}_{n=0}^{\infty}$ of Laurent polynomials can be obtained satisfying, for all $n, m \geq 0$:

1. $\psi_n(z) \in \mathcal{L}_n \setminus \mathcal{L}_{n-1}, \qquad \psi_0(z) \equiv 1,$
2. $\psi_n(z)$ has a real positive coefficient for the power $\begin{cases} z^{n-p_n} & \text{if } s_n = 0 \\ z^{-p_n} & \text{if } s_n = 1, \end{cases}$
3. $\langle \psi_n(z), \psi_m(z) \rangle = \begin{cases} 0 & \text{if } n \neq m \\ 1 & \text{if } n = m. \end{cases}$

This sequence will be called a *sequence of orthonormal Laurent polynomials for the measure* μ *and the generating sequence* $\{p_n\}_{n=0}^{\infty}$.

Let us illustrate these ideas with three examples.

Example 1. Consider the sequence of monomials given by (9) and the monomial $1 = z^0$ in the position 0. Then, the construction of the sequence $\{s_n\}_{n=1}^{\infty}$ is nothing but to take $s_n = 0$ if the nth monomial in (9) has a positive exponent and $s_n = 1$ if it is negative, whereas p_n counts the number of negative monomials positioned up to n. Hence, $\{s_n\}_{n=1}^{\infty} = \{1, 0, 1, 0, 0, 1, 1, 0, 0, \ldots\}$ and $\{p_n\}_{n=0}^{\infty} = \{0, 1, 1, 2, 2, 2, 3, 4, 4, 4, \ldots\}$.

Example 2. If $s_k = 0$ for all $k \geq 1$, then $\mathcal{L}_n$ is the space of ordinary polynomials of degree at most n. In this case the Gram–Schmidt orthogonalization process is applied to the sequence of monomials $\{1, z, z^2, z^3, \ldots\}$ and the resulting orthonormal Laurent polynomials $\psi_n(z)$ are just the well-known orthonormal *Szegő polynomials* $\varphi_n(z)$; see e.g. [30].

Example 3. If $s_k = k+1 \bmod 2$ for all $k \geq 1$, then the Gram–Schmidt orthogonalization process is applied to the sequence $\{1, z, z^{-1}, z^2, z^{-2}, \ldots\}$, where the monomials z^k and z^{-k} occur in an alternating way. The resulting sequence $\{\psi_n(z)\}_{n=0}^{\infty}$ was firstly considered by Thron in [31] and it is called the *CMV basis* in [28]. The CMV basis can actually be expressed in terms of the Szegő polynomials as (see e.g. [7,11,28,31,33])

$$\varphi_0(z), \varphi_1(z), z^{-1}\varphi_2^*(z), z^{-1}\varphi_3(z), z^{-2}\varphi_4^*(z), z^{-2}\varphi_5(z), \ldots.$$

In the general case, one has the following result.

Lemma 4 (*Cruz-Barroso et al. [12]; See also Watkins [33]*). *The family $\{\psi_n(z)\}_{n=0}^{\infty}$ is the sequence of orthonormal Laurent polynomials on the unit circle for a measure μ and the ordering induced by the generating sequence $\{p_n\}_{n=0}^{\infty}$, if and only if,*

$$\psi_n(z) = \begin{cases} z^{-p_n}\varphi_n(z) & \text{if } s_n = 0, \\ z^{-p_n}\varphi_n^*(z) & \text{if } s_n = 1, \end{cases} \tag{13}$$

$\{\varphi_n(z)\}_{n=0}^{\infty}$ being the sequence of orthonormal Szegő polynomials for μ. □

Lemma 4 shows that the orthonormal Laurent polynomials $\{\psi_n(z)\}_n$ are very closely related to the usual Szegö polynomials $\{\varphi_n(z)\}_n$ and their duals, and this for *any* choice of the generating sequence $\{p_n\}_{n=0}^{\infty}$. We will need this result in what follows.

2.2. The main result

In this subsection we state and prove the main result of this paper. Let $\{\psi_n(z)\}_{n=0}^{\infty}$ be the sequence of orthonormal Laurent polynomials on the unit circle for the measure μ and the ordering induced by the generating sequence $\{p_n\}_{n=0}^{\infty}$. To distinguish them from the other orthonormal sequences to be constructed in this section, we will equip these Laurent polynomials with a superscript: $\psi_n^{(0)}(z) := \psi_n(z)$. We will also find it convenient to use the vectorial notation $\boldsymbol{\psi}^{(0)}(z) := (\psi_n^{(0)}(z))_{n=0}^{\infty}$. Thus, $\boldsymbol{\psi}^{(0)}$ is an infinite-dimensional vector whose nth component is the nth orthonormal Laurent polynomial $\psi_n^{(0)}$ $(n \geq 0)$.

Let M denote the operator of multiplication by z on the space of quadratically integrable functions with respect to the inner product (11). Thus M is defined by the action

$$M : f(z) \mapsto zf(z), \quad f \in L_2^{\mu}(\mathbb{T}).$$

Since we are working on the unit circle $\mathbb{T}$, the operator M is actually *unitary*.

Recall the notation $\mathcal{L}$ for the closure in $L_2^{\mu}(\mathbb{T})$ of the subspace generated by $\boldsymbol{\psi}^{(0)}(z)$. We distinguish between three cases:

1. If $\lim_{n\to\infty} p_n = \lim_{n\to\infty}(n - p_n) = \infty$, then $\mathcal{L} = L_2^{\mu}(\mathbb{T})$. The sequence of orthonormal Laurent polynomials $\boldsymbol{\psi}^{(0)}$ forms then a basis for $L_2^{\mu}(\mathbb{T})$ and the matrix representation of M with respect to this basis is an infinite unitary matrix $\mathcal{S}$, i.e., both the rows and columns of this matrix are orthonormal.
2. If $\lim_{n\to\infty} p_n < \infty$, then the sequence of orthonormal Laurent polynomials $\boldsymbol{\psi}^{(0)}$ can be non-complete, but in any way it will still generate a subspace of $L_2^{\mu}(\mathbb{T})$ which is invariant under the application of the operator M. We can then define the operator $M \restriction \mathcal{L}$, which is the restriction of a unitary operator to an invariant subspace and hence is isometric. The matrix representation of this operator with respect to the basis $\boldsymbol{\psi}^{(0)}$ of $\mathcal{L}$ is now an infinite isometric matrix $\mathcal{S}$, i.e., the columns of this matrix are orthonormal. In fact, it is known that the sequence $\boldsymbol{\psi}^{(0)}$ is complete in $L_2^{\mu}(\mathbb{T})$, if and only if the so-called *Szegö condition* fails, i.e., if $\sum_{j=0}^{\infty}|\alpha_j|^2 = \infty$. In that case the matrix $\mathcal{S}$ is actually *unitary* since $M \restriction \mathcal{L} = M$.
3. If $\lim_{n\to\infty}(n - p_n) < \infty$,[3] then the sequence of orthonormal Laurent polynomials $\boldsymbol{\psi}^{(0)}$ can be non-complete, and in that case it generates a subspace of $L_2^{\mu}(\mathbb{T})$ which is *not* invariant under

[3] We thank the referee for pointing our attention to this case, and for providing us with the modifications that have to be made for it.

the application of the operator M. However, we can now still consider the operator $PM \restriction \mathcal{L}$ where P is the orthogonal projection operator of $L_2^\mu(\mathbb{T})$ onto $\mathcal{L}$. The matrix representation $\mathcal{S}$ of this operator with respect to the basis $\boldsymbol{\psi}^{(0)}$ of $\mathcal{L}$ is now not necessarily unitary neither isometric. Actually, it holds that the *rows* of this matrix are orthonormal. This follows by noticing that the transpose of the matrix $\mathcal{S}$ occurs as a matrix representation in the previous case and hence is isometric.

Note that in each of the above three cases, the infinite matrix $\mathcal{S}$ has its entries given by

$$\begin{aligned}\mathcal{S} &= [\langle \psi_i^{(0)}(z), z\psi_j^{(0)}(z)\rangle]_{i,j=0}^{\infty} \\ &=: \langle \boldsymbol{\psi}^{(0)}(z), z\boldsymbol{\psi}^{(0)}(z)\rangle, \end{aligned} \tag{14}$$

where the inner product is defined in (11). Here the expression on the second line should be regarded as a compact vectorial notation of the line above. Now we are in position to prove the following result. We will do this by using a modification of an argument of Simon [28, third proof of theorem 10.1] for the isometric Hessenberg case. The main ingredient of the proof will be the well-known *Szegő recursion*, expressed in the form (see e.g. [30])

$$\begin{bmatrix} z\varphi_k(z) \\ \varphi_{k+1}^*(z) \end{bmatrix} = \begin{bmatrix} \overline{\alpha_k} & \rho_k \\ \rho_k & -\alpha_k \end{bmatrix} \begin{bmatrix} \varphi_k^*(z) \\ \varphi_{k+1}(z) \end{bmatrix}, \tag{15}$$

where $\varphi_k(z)$ and $\varphi_k^*(z)$ denote the orthonormal Szegő polynomial of degree k and its dual respectively. Note that the coefficient matrix in (15) is nothing but the non-trivial part (5) of the Givens transformation $G_{k,k+1}$.

Theorem 5. *Let $\{\psi_n(z)\}_{n=0}^{\infty}$ be the sequence of orthonormal Laurent polynomials on the unit circle for a measure μ and the ordering induced by the generating sequence $\{p_n\}_{n=0}^{\infty}$. Then the matrix $\mathcal{S}$ in (14) can be factored into a snake-shaped matrix factorization $\mathcal{S} = \mathcal{S}^{(\infty)}$, constructed by the recipe given in Section 1.3. The factorization must be understood in the sense that the principal $n \times n$ submatrices of $\mathcal{S}^{(n-1)}$ and $\mathcal{S}$ coincide for all n.*

Proof. We will construct a sequence of intermediary bases $\boldsymbol{\psi}^{(k)}$ for the subspace $\mathcal{L} = \text{span}\{\psi_j^{(0)}\}_{j=0}^{\infty}$, $k \geq 1$, in such a way that for each k, there exists an index $l \in \{0, 1, \ldots, k-1\}$ such that $\boldsymbol{\psi}^{(k)}$ is the same as $\boldsymbol{\psi}^{(l)}$, except for a change in the $(k-1)$th and kth components. These intermediary bases will serve to factorize the matrix $\mathcal{S}$. For example, note that (14) can be rewritten as

$$\mathcal{S} = \langle \boldsymbol{\psi}^{(0)}(z), z\boldsymbol{\psi}^{(0)}(z)\rangle \tag{16}$$

$$= \langle \boldsymbol{\psi}^{(0)}(z), \boldsymbol{\psi}^{(1)}(z)\rangle \cdot \langle \boldsymbol{\psi}^{(1)}(z), z\boldsymbol{\psi}^{(0)}(z)\rangle, \tag{17}$$

for any choice of the basis $\boldsymbol{\psi}^{(1)}$ of $\mathcal{L}$. Indeed, the jth column of the matrix (17) is obtained by expressing the orthogonal projection on $\mathcal{L}$ of the function $z\psi_j^{(0)}(z)$ in terms of the basis $\boldsymbol{\psi}^{(1)}(z)$, which is then in its turn expressed in terms of the basis $\boldsymbol{\psi}^{(0)}(z)$. Obviously this gives the same result as directly expressing the orthogonal projection on $\mathcal{L}$ of $z\psi_j^{(0)}(z)$ in terms of the basis $\boldsymbol{\psi}^{(0)}(z)$, i.e., it equals the jth column of (16) (we recall again our convention with the notation: $j \geq 0$).

Note that instead of (17) we could also have written a slightly modified version of it:

$$\begin{aligned}\mathcal{S} &= \langle \boldsymbol{\psi}^{(0)}(z), z\boldsymbol{\psi}^{(0)}(z)\rangle \\ &= \langle \boldsymbol{\psi}^{(0)}(z), z\boldsymbol{\psi}^{(1)}(z)\rangle \cdot \langle z\boldsymbol{\psi}^{(1)}(z), z\boldsymbol{\psi}^{(0)}(z)\rangle \\ &= \langle \boldsymbol{\psi}^{(0)}(z), z\boldsymbol{\psi}^{(1)}(z)\rangle \cdot \langle \boldsymbol{\psi}^{(1)}(z), \boldsymbol{\psi}^{(0)}(z)\rangle \end{aligned} \tag{18}$$

for any choice of the basis $\boldsymbol{\psi}^{(1)}$ of $\mathcal{L}$ and where we have used the general fact that $\langle zf(z), zg(z)\rangle = \langle f(z), g(z)\rangle$ for any functions $f, g : \mathbb{T} \to \mathbb{C}$, which follows from (11) and the fact that $z \in \mathbb{T}$. The choice between (17) and (18) will depend on the fact whether $s_1 = 0$ or $s_1 = 1$, respectively; see further.

The point will now be to make a good choice for the intermediary bases $\boldsymbol{\psi}^{(k)}$. For example, the 'good' choice for $\boldsymbol{\psi}^{(1)}$ will be the one for which one of the factors in (17) (or (18)) equals the Givens transformation $G_{0,1}$, while the other factor is of the form

$$\begin{bmatrix} 1 & 0 \\ 0 & * \end{bmatrix},$$

where $*$ denotes an irrelevant submatrix (which is actually of infinite dimension). Explicitly, the basis $\boldsymbol{\psi}^{(1)}$ is given by $\boldsymbol{\psi}_0^{(1)} = z^{1-2s_1}$, $\boldsymbol{\psi}_1^{(1)} = z^{-s_1}[z^{s_1}\boldsymbol{\psi}_1^{(0)}]^*$ and $\boldsymbol{\psi}_k^{(1)} = \boldsymbol{\psi}_k^{(0)}$ for all $k \geq 2$. Repeating this idea inductively for all subsequent bases $\boldsymbol{\psi}^{(k)}$ will ultimately lead to the decomposition of $\mathcal{S}$ as an infinite product of Givens transformations.

Let us now formalize these ideas. We work with the induction hypothesis that after the kth step, $k \geq 0$,[4] we have decomposed the matrix $\mathcal{S}$ as

$$\begin{aligned}\mathcal{S} &= S^{(k-1)}X^{(k)}, \quad \text{if } s_k = 0, \\ \mathcal{S} &= X^{(k)}S^{(k-1)}, \quad \text{if } s_k = 1,\end{aligned} \tag{19}$$

where $\mathcal{S}^{(k-1)}$ is the $(k-1)$th iterate matrix of the snake-shaped matrix factorization $\mathcal{S}^{(\infty)}$ (cf. the construction in Section 1.3), while $X^{(k)}$ equals the identity matrix in its first k rows and columns, i.e.,

$$X^{(k)} = \begin{bmatrix} I_k & 0 \\ 0 & * \end{bmatrix}, \tag{20}$$

with I_k the identity matrix of size k. We also assume by induction that

$$X^{(k)} = \langle \boldsymbol{\psi}^{(l)}(z), z\boldsymbol{\psi}^{(m)}(z)\rangle, \tag{21}$$

where l, m are certain indices in $\{0, 1, \dots, k\}$ with at least one of them equal to k (we could actually give explicit expressions for l, m but will not need these in what follows). Note that by combining the hypotheses (20) and (21), we deduce that $\psi_i^{(l)}(z) = z\psi_i^{(m)}(z)$ for all $i \in \{0, 1, \dots, k-1\}$. In addition, we have the following induction hypothesis on the kth components of $\boldsymbol{\psi}^{(l)}$ and $\boldsymbol{\psi}^{(m)}$:

$$\begin{aligned}\psi_k^{(l)}(z) &= z^{-p_k}\varphi_k^*(z), \\ \psi_k^{(m)}(z) &= z^{-p_k}\varphi_k(z),\end{aligned} \tag{22}$$

where $\varphi_k(z)$ denotes the orthonormal Szegő polynomial of degree k.

[4] This procedure also works for $k = 0$ provided that we set $\mathcal{S}^{(-1)} := I$ and $s_0 = 0$ or $s_0 = 1$, since either choice will give the same result.

Identities (19) and (20) imply the coincidence of the principal $k \times k$ submatrices of $S^{(k-1)}$ and S, as the theorem states. Thus, to prove the theorem we simply must show that, given all the above induction hypotheses, we can now come to the induction step $k \mapsto k+1$. To this end, we should try to peel off a new Givens transformation $G_{k,k+1}$ from the matrix $\mathcal{S}$. Assume that the first k intermediary bases $\boldsymbol{\psi}^{(1)}, \ldots, \boldsymbol{\psi}^{(k)}$ of $\mathcal{L}$ have already been constructed. We want to define the next intermediary basis $\boldsymbol{\psi}^{(k+1)}$. We distinguish between two cases:

1. If $s_{k+1} = 0$, we define $\boldsymbol{\psi}^{(k+1)}$ to be the same as $\boldsymbol{\psi}^{(l)}$, except for its kth and $(k+1)$th components. More precisely, we set

$$\begin{bmatrix} \psi_k^{(k+1)}(z) \\ \psi_{k+1}^{(k+1)}(z) \end{bmatrix} := \tilde{G}_{k,k+1} \begin{bmatrix} \psi_k^{(l)}(z) \\ \psi_{k+1}^{(l)}(z) \end{bmatrix} \tag{23}$$

$$= \tilde{G}_{k,k+1} \cdot z^{-p_{k+1}} \begin{bmatrix} \varphi_k^*(z) \\ \varphi_{k+1}(z) \end{bmatrix} \tag{24}$$

$$= z^{-p_{k+1}} \begin{bmatrix} z\varphi_k(z) \\ \varphi_{k+1}^*(z) \end{bmatrix}. \tag{25}$$

Here the second equality follows from the first lines of (22) and (13) (recall that the $(k+1)$th component of $\boldsymbol{\psi}^{(l)}$ has not been changed yet with respect to $\boldsymbol{\psi}^{(0)}$), and from the fact that $p_{k+1} = p_k$ by assumption. On the other hand, the third equality is nothing but the Szegő recursion (15).

Then, we can factorize (21) as

$$\begin{aligned} X^{(k)} &= \langle \boldsymbol{\psi}^{(l)}(z), z\boldsymbol{\psi}^{(m)}(z) \rangle \\ &= \langle \boldsymbol{\psi}^{(l)}(z), \boldsymbol{\psi}^{(k+1)}(z) \rangle \cdot \langle \boldsymbol{\psi}^{(k+1)}(z), z\boldsymbol{\psi}^{(m)}(z) \rangle \\ &= G_{k,k+1} \cdot \langle \boldsymbol{\psi}^{(k+1)}(z), z\boldsymbol{\psi}^{(m)}(z) \rangle \\ &=: G_{k,k+1} X^{(k+1)}, \end{aligned} \tag{26}$$

where the third equality follows from (23), and where the matrix $X^{(k+1)}$ in the fourth equality now equals the identity matrix in its first $k+1$ rows and columns. The latter follows by the induction hypothesis for the first k rows and columns (rows and columns 0 to $k-1$), and from the fact that, by the first line of (25) and the second line of (22), we have

$$\psi_k^{(k+1)}(z) = z \cdot z^{-p_{k+1}} \varphi_k(z) = z\psi_k^{(m)}(z),$$

implying that also the kth column of the matrix $X^{(k+1)}$ has all its entries equal to zero, except for the diagonal entry which equals one. From the fact that the kth row of the matrix $X^{(k+1)}$ is a vector with norm at most one and with one of its entries equal to one, it then follows that also the kth *row* has all its entries equal to zero, except for the diagonal entry. We can then replace the index l by its new value $k+1$. We have already checked that the induction hypotheses (20) and (21) are inherited in this way as $k \mapsto k+1$. Also the hypothesis (22) can be easily checked to remain valid in this way, by virtue of the second line of (25) and the first line of (13). Finally, we have to check that (19) remains also valid. To prove this, we use (19) and (26), the construction of $\mathcal{S}^{(k)}$ in Section 1.3 and we distinguish between two cases:

(a) If $s_k = 0$ then

$$\begin{aligned} \mathcal{S} &= \mathcal{S}^{(k-1)} X^{(k)} \\ &= \mathcal{S}^{(k-1)} G_{k,k+1} X^{(k+1)} \\ &=: \mathcal{S}^{(k)} X^{(k+1)}. \end{aligned}$$

(b) If $s_k = 1$ then

$$\begin{aligned}\mathcal{S} &= X^{(k)}\mathcal{S}^{(k-1)}\\ &= G_{k,k+1}X^{(k+1)}\mathcal{S}^{(k-1)}\\ &= G_{k,k+1}\mathcal{S}^{(k-1)}X^{(k+1)}\\ &=: \mathcal{S}^{(k)}X^{(k+1)},\end{aligned}$$

where we have used the commutativity of $\mathcal{S}^{(k-1)}$ and $X^{(k+1)}$ since these matrices have a complementary zero pattern (the former equals the identity matrix except for its first $(k+1)\times(k+1)$ block, while the latter is precisely the identity matrix there, cf. (20)).

2. If $s_{k+1} = 1$, we define $\boldsymbol{\psi}^{(k+1)}$ to be the same as $\boldsymbol{\psi}^{(m)}$, except for its kth and $(k+1)$th components. More precisely, we set

$$\begin{bmatrix}\psi_k^{(k+1)}(z)\\ \psi_{k+1}^{(k+1)}(z)\end{bmatrix} := \tilde{G}_{k,k+1}^{-1}\begin{bmatrix}\psi_k^{(m)}(z)\\ \psi_{k+1}^{(m)}(z)\end{bmatrix} \tag{27}$$

$$= \tilde{G}_{k,k+1}^{-1}\cdot z^{-p_{k+1}}\begin{bmatrix}z\varphi_k(z)\\ \varphi_{k+1}^*(z)\end{bmatrix} \tag{28}$$

$$= z^{-p_{k+1}}\begin{bmatrix}\varphi_k^*(z)\\ \varphi_{k+1}(z)\end{bmatrix}, \tag{29}$$

where we have used the second lines of (22) and (13) (recall that the $(k+1)$th component of $\boldsymbol{\psi}^{(m)}$ has not been changed yet with respect to $\boldsymbol{\psi}^{(0)}$), the fact that $p_{k+1} = p_k + 1$ by assumption and the Szegő recursion (15).

Then, we can factorize (21) as

$$\begin{aligned}X^{(k)} &= \langle \boldsymbol{\psi}^{(l)}(z), z\boldsymbol{\psi}^{(m)}(z)\rangle\\ &= \langle \boldsymbol{\psi}^{(l)}(z), z\boldsymbol{\psi}^{(k+1)}(z)\rangle\cdot\langle z\boldsymbol{\psi}^{(k+1)}(z), z\boldsymbol{\psi}^{(m)}(z)\rangle\\ &= \langle \boldsymbol{\psi}^{(l)}(z), z\boldsymbol{\psi}^{(k+1)}(z)\rangle\cdot\langle \boldsymbol{\psi}^{(k+1)}(z), \boldsymbol{\psi}^{(m)}(z)\rangle\\ &= \langle \boldsymbol{\psi}^{(l)}(z), z\boldsymbol{\psi}^{(k+1)}(z)\rangle\cdot G_{k,k+1}\\ &=: X^{(k+1)}G_{k,k+1},\end{aligned}$$

where the fourth step follows from (27).

It is easy to check again that $X^{(k+1)}$ equals the identity matrix in its first $k+1$ rows and columns by using the induction hypothesis for the first k rows and columns $(0, 1, \ldots, k-1)$ and from the first lines of (29) and (22) for the kth row and column.

We can then replace the index m by its new value $k+1$. If follows from the above discussion that the induction hypotheses (20) and (21) are inherited in this way as $k \mapsto k+1$. Also the hypothesis (22) goes through, by virtue of the second lines of (13) and (29). Finally, the proof that also (19) goes through can be proven in a completely similar way as in the previous case.

We have now completely established the induction hypothesis $k \mapsto k+1$, hereby ending the proof of Theorem 5. □

3. Entry-wise expansion of a snake-shaped matrix factorization

In this section we discuss the entry-wise expansion of a snake-shaped matrix factorization $\mathcal{S}$, hereby generalizing the expansions in (1) and (2).

3.1. Graphical rule for the entry-wise expansion of $\mathcal{S}$

First we will present a graphical rule for predicting both the position and the form of the non-zero entries of a snake-shaped matrix factorization $\mathcal{S}$.

We will illustrate the ideas for the matrix $\mathcal{S}$ given by (10) and Fig. 3. A straightforward computation shows that the full expansion of this matrix $\mathcal{S}$ is given by (compare with [6], Example 4.5)

$$\begin{pmatrix}
\overline{\alpha_0} & \rho_0 & 0 & 0 & 0 & 0 & 0 & 0\cdots \\
\rho_0\overline{\alpha_1} & -\alpha_0\overline{\alpha_1} & \rho_1\overline{\alpha_2} & \rho_1\rho_2 & 0 & 0 & 0 & 0\cdots \\
\rho_0\rho_1 & -\alpha_0\rho_1 & -\alpha_1\overline{\alpha_2} & -\alpha_1\rho_2 & 0 & 0 & 0 & 0\cdots \\
0 & 0 & \rho_2\overline{\alpha_3} & -\alpha_2\overline{\alpha_3} & \rho_3\overline{\alpha_4} & \rho_3\rho_4\overline{\alpha_5} & \rho_3\rho_4\rho_5 & 0\cdots \\
0 & 0 & \rho_2\rho_3 & -\alpha_2\rho_3 & -\alpha_3\overline{\alpha_4} & -\alpha_3\rho_4\overline{\alpha_5} & -\alpha_3\rho_4\rho_5 & 0\cdots \\
0 & 0 & 0 & 0 & \rho_4 & -\alpha_4\overline{\alpha_5} & -\alpha_4\rho_5 & 0\cdots \\
0 & 0 & 0 & 0 & 0 & \rho_5\overline{\alpha_6} & -\alpha_5\overline{\alpha_6} & \rho_6\cdots \\
0 & 0 & 0 & 0 & 0 & \rho_5\rho_6\overline{\alpha_7} & -\alpha_5\rho_6\overline{\alpha_7} & -\alpha_6\overline{\alpha_7}\cdots \\
0 & 0 & 0 & 0 & 0 & \rho_5\rho_6\rho_7 & -\alpha_5\rho_6\rho_7 & -\alpha_6\rho_7\cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}. \tag{30}$$

Now the attentive reader will notice that the zero pattern of this matrix $\mathcal{S}$ has some similarity with the shape of its underlying snake as shown in Fig. 3. Actually, we claim that the (i, j) entry of the matrix $\mathcal{S}$ can be obtained from the following recipe (the 'E' stands for 'entry-wise'):

E1. Draw the snake underlying the matrix $\mathcal{S}$ (cf. Fig. 3);
E2. Place a right-pointing arrow on the left of the snake at height i;
E3. Place a left-pointing arrow on the right of the snake at height j;
E4. Draw the *path* on the snake induced between these two arrows;
E5. If the path moves monotonically from left to right, then the (i, j) entry of $\mathcal{S}$ equals a product of entries of the encountered Givens transformations

$$\tilde{G}_{k,k+1} = \begin{bmatrix} \overline{\alpha_k} & \rho_k \\ \rho_k & -\alpha_k \end{bmatrix} \tag{31}$$

on the path (see Step E5′ below for a specification of this rule);
E6. If the path does *not* move monotonically from left to right, then the (i, j) entry of S equals zero.

Let us illustrate this recipe for the (7, 5) entry of the matrix $\mathcal{S}$ (recall that we label the rows and columns of this matrix starting from the index 0). The recipe is shown for this case in Fig. 4.

Let us comment on Fig. 4. Fig. 4(a) shows the snake shape of the matrix $\mathcal{S}$ (compare with Fig. 3), corresponding to Step E1 in the above recipe. Fig. 4(b) shows the arrows on the left and on the right of the snake at height 7 and 5, respectively, corresponding to Steps E2 and E3. The path on the snake induced between these two arrows is shown in Fig. 4(c), corresponding to Step E4. Note that this path moves monotonically from left to right and passes through the Givens transformations $G_{7,8}$, $G_{6,7}$ and $G_{5,6}$. From Step E5 it then follows that the (7, 5) entry of the matrix $\mathcal{S}$ is a product of entries of these three Givens transformations. Actually, it equals $\rho_5\rho_6\overline{\alpha_7}$ (compare with (30)).

As a second example, let us consider the (7, 4) entry of the matrix S. The recipe is illustrated for this case in Fig. 5.

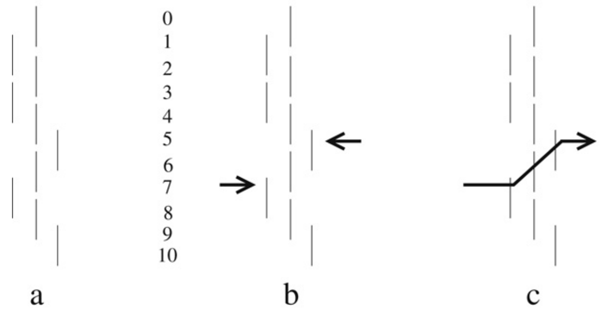

Fig. 4. The figure shows (a) the snake shape underlying the matrix $\mathcal{S}$ in Eq. (10), (b) the arrows on the left and on the right of the snake at height 7 and 5, respectively, and (c) the path on the snake induced between these two arrows. From this information, the value of the (7, 5) entry of the matrix $\mathcal{S}$ can be determined.

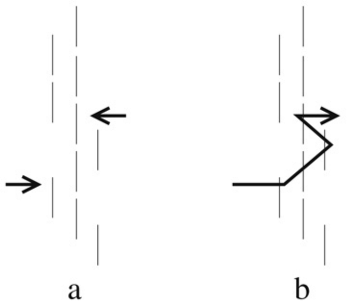

Fig. 5. For the matrix $\mathcal{S}$ in Eq. (3), the figure shows (a) the arrows on the left and on the right of the snake at height 7 and 4, respectively, and (b) the path on the snake induced between these two arrows. Since now the path does *not* move monotonically from left to right, it follows that the (7, 4) entry of $\mathcal{S}$ equals zero. This corresponds again with (30).

Note that in the above example concerning the (7, 5) entry of the matrix $\mathcal{S}$, we noticed that this entry equals the product of the (complex conjugate of) the *Schur parameter* $\overline{\alpha_7}$, on the one hand, and the *complementary Schur parameters* ρ_6, ρ_5, on the other hand. To complete our description, let us now state an *a priori* rule to determine which of the four entries in (31) each Givens transformation $G_{k,k+1}$ on the path in Step E5 contributes.

Let us explain this rule for the first Givens transformation $G_{7,8}$ through which the path in Fig. 4(c) passes (note that $G_{7,8}$ corresponds to the bottom leftmost line segment on the path in Fig. 4(c)). First, we will determine the *row index* of the entry contributed by $G_{7,8}$. To this end, imagine that we are in the line segment corresponding to $G_{7,8}$ and that we move *leftwards* on the path. It is then seen from Fig. 4(c) that we leave this line segment through its topmost index; hence we claim that the sought entry of $\tilde{G}_{7,8}$ will be in its topmost row.

Next, to find the *column index* of the entry contributed by $\tilde{G}_{7,8}$, imagine again that we start in the line segment corresponding to $G_{7,8}$ but move this time *rightwards* on the path. Since the path in Fig. 4(c) proceeds upwards from left to right, we move then to the position of smaller indices. Hence the sought entry of $\tilde{G}_{7,8}$ will be in its column with the smallest index, which is column 0. We conclude that the sought entry of $\tilde{G}_{7,8}$ lies in the (0, 0) position of (31); this gives us $\overline{\alpha_7}$.

The entries contributed by $G_{6,7}$ and $G_{5,6}$ can be found in a similar way. The reader can check that in both cases, the relevant entries of $G_{6,7}$, $G_{5,6}$ are positioned in the (1, 0) entry of (31).

To summarize these ideas, let us introduce some notations. Denote with $G_{r,r+1}$ and $G_{t,t+1}$ the two *outermost* line segments of the path in Step E5. Note that $r \in \{i-1, i\}$ and $t \in \{j-1, j\}$, with the precise value of r and t depending on the shape of the snake. For the example of Fig. 4(c), we have $r = i = 7$ and $t = j = 5$.

Denote with $\mathcal{K}$ the set of indices k of the *innermost* Givens transformations $G_{k,k+1}$ on the path. Explicitly, $\mathcal{K}$ equals $\{r+1, \ldots, t-1\}$ if $r < t$ and $\{t+1, \ldots, r-1\}$ if $r > t$ (it is understood that $\mathcal{K} = \emptyset$ when $|r-t| = 1$).

It is easily seen from the above graphical rule that the $\{G_{k,k+1}\}_{k\in\mathcal{K}}$ in Step E5 always contribute their *complementary Schur parameter* ρ_k, while $G_{r,r+1}$ and $G_{t,t+1}$ can contribute each of their entries. In fact, we state the following specification of Step E5.

E5′. Under the assumptions of Step E5, and using the above notations, the (i, j) entry of the matrix $\mathcal{S}$ equals

$$x_r \cdot \left(\prod_{k\in\mathcal{K}} \rho_k\right) \cdot y_t,$$

where $x_r \in \{\overline{\alpha_r}, \rho_r, -\alpha_r\}$ and $y_t \in \{\overline{\alpha_t}, \rho_t, -\alpha_t\}$ are the entries of $\tilde{G}_{r,r+1}$ and $\tilde{G}_{t,t+1}$ which can be found as described in the paragraphs above[5]: it suffices each time to imagine that we are in the line segment corresponding to the current Givens transformation, and then imagine moving leftwards or rightwards on the path, to obtain the row and the column index in (31), respectively.

3.2. *Proof of the graphical rule*

The proof that the recipe in Steps E1–E6, E5′ leads to the correct form of the (i, j) entry of $\mathcal{S}$ follows by just expanding the matrix $\mathcal{S}$ in an appropriate way. Let us sketch here the main steps of the proof.

Proof. Throughout the proof, the Givens transformations under the $\prod$-symbol are understood to be multiplied in the order described in Section 1.3. We will assume for definiteness that either $i < j$ or $i = j$ and $r < t$. Consider the given snake-shaped matrix factorization $\mathcal{S} = \prod_{k=0}^{\infty} G_{k,k+1}$. Define the 'sub-snake'

$$\mathcal{S}_{i,j} := \prod_{k=i-1}^{j} G_{k,k+1}. \tag{32}$$

It is clear that the (i, j) entry of $\mathcal{S}$ depends only on the sub-snake $\mathcal{S}_{i,j}$. This follows since the other Givens transformations can be considered as operations on rows and columns $\{1, 2, \ldots, i-1\} \cup \{j+1, j+2, \ldots\}$ of $\mathcal{S}_{i,j}$; hence indeed they cannot influence the (i, j) entry of $\mathcal{S}_{i,j}$.

Assume now that $s_l = 1$ for some $l \in \{i, \ldots, j\}$. This means that the line segment $G_{l,l+1}$ is positioned to the left of $G_{l-1,l}$. We can then factor (32) as

$$\left(\prod_{k=l}^{j} G_{k,k+1}\right) \cdot \left(\prod_{k=i-1}^{l-1} G_{k,k+1}\right). \tag{33}$$

[5] Explicitly, x_r is the $(i-r, b)$th entry of $\tilde{G}_{r,r+1}$ and y_t is the $(1-b, j-t)$th entry of $\tilde{G}_{t,t+1}$, where the Boolean b is defined by $b = 0$ if $r > t$ or $b = 1$ if $r < t$.

We distinguish between three cases:

- Suppose that $l \in \{i+1, \ldots, j-1\}$. The leftmost factor in (33) can be considered as a row operation acting on rows $l, \ldots, j+1$ of the rightmost factor in (33). By assumption, these row indices are all strictly larger than i; hence this factor cannot influence the (i, j) entry of $\mathcal{S}_{i,j}$. Similarly, the rightmost factor in (33) acts on columns $i-1, \ldots, l$, which by assumption are all strictly smaller than j. We conclude that the (i, j) entry can be influenced by *none* of the factors $G_{k,k+1}$ in (33), and hence it simply equals the (i, j) entry of the identity matrix, i.e., it equals zero. This proves the conclusion in Step E6.
- Suppose that $l = i$. In contrast to the previous case, we can now only conclude that the rightmost factor $G_{i-1,i}$ in (33) can be removed from further consideration. This corresponds to the fact that r equals i (and not $i-1$) in this case.
- Suppose that $l = j$. Similarly as in the previous case, we can then conclude that the leftmost factor $G_{j,j+1}$ in (33) can be removed from further consideration. This corresponds to the fact that t equals $j-1$ (and not j) in this case.

Getting rid of all the redundant factors $G_{k,k+1}$ as described above, we are left with either the identity matrix or with a sequence of Givens transformations following a unitary Hessenberg shape (cf. Fig. 1). The relevant entries of this matrix can be computed using a straightforward calculation and are easily seen to correspond to the given rules in Steps E5 and E5$'$ (compare with (2)). We omit further details. $\square$

3.3. Some corollaries

A first corollary is the following.

Corollary 6 (*Upper and Lower Bandwidth of $\mathcal{S}$*). *The upper bandwidth of the snake-shaped matrix factorization $\mathcal{S}$ equals the length of the longest sub-snake of $\mathcal{S}$ whose line segments are linearly aligned in the top left-bottom right order (cf. Fig. 1). Similarly, the lower bandwidth of $\mathcal{S}$ equals the length of the longest sub-snake of $\mathcal{S}$ whose line segments are linearly aligned in the top right-bottom left order.*

It follows from Corollary 6 that the unitary five-diagonal matrices $\mathcal{C}$ have the smallest bandwidth of all snake-shaped matrix factorizations $\mathcal{S}$; they have in fact bandwidth 2 in both their lower and upper triangular part and hence are five-diagonal.

A related result on the minimality of the matrix $\mathcal{C}$ is the fact [8] that any infinite unitary matrix A having lower bandwidth 1 and finite upper bandwidth n is 'trivial' in the sense that A can be decomposed as a direct sum of matrices of size at most $n+1$. This result can be shown using only some basic linear algebra by noting that under the above conditions on the matrix A, this matrix is isometric Hessenberg and hence allows a factorization of the form (4). The condition on the upper bandwidth of A then easily implies that from each tuple of $n+1$ subsequent Givens transformations $G_{k,k+1}$ in (4), there must be at least one for which $G_{k,k+1}$ has vanishing off-diagonal elements; we omit further details.

A second corollary of the above results can be easily proven from (14) and Lemma 4. Here, the elements of the matrix $\mathcal{S}$ given by (14) are expressed in terms of the inner product (11) and the orthonormal Szegő polynomials (see also Theorem 4.1 in [6]).

Corollary 7. *By introducing the notation*

$$f_i = \begin{cases} \varphi_i(z) & \text{if } s_i = 0, \\ \varphi_i^*(z) & \text{if } s_i = 1, \end{cases}$$

then the entries of the snake-shaped matrix factorization $\mathcal{S} = \left(\eta_{i,j}\right)_{i,j\geq 0}$ *are given for all* $i \geq 0$ *and* $k \geq 1$ *by* $\eta_{i,i} = \langle f_i, zf_i\rangle$ *and by*

$$\eta_{i+k,i} = \begin{cases} \langle f_{i+k}, z^{k+s_{i+k}} f_i\rangle & \text{if } s_{i+1} = \cdots = s_{i+k-1} = 1, \\ 0 & \text{other case,} \end{cases}$$
$$\eta_{i,i+k} = \begin{cases} \langle f_i, z^{1-s_{i+k}} f_{i+k}\rangle & \text{if } s_{i+1} = \cdots = s_{i+k-1} = 0, \\ 0 & \text{other case,} \end{cases}$$

where when $k = 1$*, the condition* $s_{i+1} = \cdots = s_{i+k-1} \in \{0, 1\}$ *is understood to be always valid.* □

As a consequence of Corollary 7 and the graphical rule, by choosing appropriate generating sequences one can easily deduce a direct proof of Propositions 1.5.8, 1.5.9 and 1.5.10 in [27].

4. Connection with Szegő quadrature formulas

In this section we describe some connections between snake-shaped matrix factorizations and Szegő quadrature formulas. The results in this section are actually known for the isometric Hessenberg and unitary five-diagonal cases, and the extension to a general snake-shaped matrix factorization $\mathcal{S}$ turns out to be rather trivial. Nevertheless, we include these results here for completeness of the paper.

Throughout this section, we shall be dealing with a fixed measure μ as described in Section 2.1 and we will be concerned with the computation of integrals of the form

$$I_\mu(f) := \int_{\mathbb{T}} f(z)\mathrm{d}\mu(z) = \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\mathrm{d}\mu(\theta),$$

by means of the so-called *Szegő quadrature formulas*. Such rules appear as the analogue on the unit circle of the Gaussian formulas when dealing with estimations of integrals supported over intervals on the real line $\mathbb{R}$. For a fixed positive integer $n \in \mathbb{N} \setminus \{0\}$, an n-point Szegő quadrature is of the form

$$I_n(f) := \sum_{j=1}^{n} \lambda_j f(z_j), \quad z_j \in \mathbb{T}, j = 1, \ldots, n, z_j \neq z_k \text{ if } j \neq k,$$

where the *nodes* $\{z_j\}_{j=1}^n$ and *weights* $\{\lambda_j\}_{j=1}^n$ are determined in such a way that the quadrature formulas are exact in subspaces of Laurent polynomials whose dimension is as high as possible. The characterizing property is that $I_n(L) = I_\mu(L)$ for all $L \in \operatorname{span}\{z^j : j = -n+1, \ldots, n-1\}$ (the optimal subspace): see e.g. [12,18,23], [20, Chapter 4].

In what follows, we will use the notations $\mathcal{H}$, $\mathcal{C}$ and $\mathcal{S}$ for the isometric Hessenberg, unitary five-diagonal and snake-shaped matrix factorization induced by the generating sequence $\{p_n\}_n$, respectively. As we have already seen, these matrices can all be factorized as $\prod_{k=0}^{\infty} G_{k,k+1}$, where the $G_{k,k+1}$ are canonically fixed by (3) and (5), but where the factors under the $\prod$-symbol may occur in a certain order (cf. Section 1.3).

We start with the following result, which seems to be essentially[6] due to Gragg. It is the unitary analogue of a well-known result for the Jacobi matrix when the measure μ is supported on the real line $\mathbb{R}$.

[6] Theorem 8 is not explicitly stated in [18], but it can be easily deduced from the results in that paper.

Theorem 8 (*Gragg [18]*). *The eigenvalues of the principal $n \times n$ submatrix of the isometric Hessenberg matrix $\mathcal{H}$ are the zeros of the nth Szegö polynomial $\varphi_n(z)$.*

Here with the principal $n \times n$ submatrix of $\mathcal{H}$ we mean the submatrix formed by rows and columns 0 up to $n-1$ of $\mathcal{H}$.

Proposition 9 (*Watkins [33], Cantero, Moral and Velázquez [7]*). *Theorem 8 also holds for the unitary five-diagonal matrix $\mathcal{C}$, i.e., the eigenvalues of the principal $n \times n$ submatrix of $\mathcal{C}$ are the zeros of the nth Szegö polynomial $\varphi_n(z)$.*

The above results hold in fact for any snake-shaped matrix factorization S; see further.

For the present discussion, a drawback of Theorem 8 and Proposition 9 is that the principal $n \times n$ submatrix of $\{\mathcal{H}, \mathcal{C}, \mathcal{S}\}$ is in general *not* unitary anymore and hence has eigenvalues *strictly* inside the unit disk. This means that these eigenvalues are not suited as nodes for the construction of an n-point Szegő quadrature formula.

The solution to the above drawback is to slightly modify the principal $n \times n$ submatrix of $\mathcal{S}$ in such a way that it becomes unitary. Its eigenvalues will then be distinct, exactly on the unit circle $\mathbb{T}$ and turn out to be precisely the required set of nodes.

To achieve this in practice, Gragg [18] and also Watkins [33] introduced the idea to redefine the $(n-1)$th Givens transformation $\tilde{G}_{n-1,n}$ by

$$\tilde{G}_{n-1,n} := \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{i\tilde{\theta}} \end{bmatrix}, \tag{34}$$

where $\theta, \tilde{\theta} \in \mathbb{R}$ denote arbitrary parameters (the second of them will actually be irrelevant for what follows).

With this new choice of $\tilde{G}_{n-1,n}$, we can 'absorb' the factors $e^{i\theta}, e^{i\tilde{\theta}}$ in the previous and next Givens transformation $G_{n-2,n-1}$ and $G_{n,n+1}$, respectively. This means that we redefine

$$\tilde{G}_{n-2,n-1} := \tilde{G}_{n-2,n-1} \cdot \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix}, \quad \text{if } s_{n-1} = 0, \tag{35}$$

while in case $s_{n-1} = 1$ we redefine $\tilde{G}_{n-2,n-1}$ by the same formula (35) but now with the factors multiplied in the reverse order. Similarly, we redefine

$$\tilde{G}_{n,n+1} := \begin{bmatrix} e^{i\tilde{\theta}} & 0 \\ 0 & 1 \end{bmatrix} \cdot \tilde{G}_{n,n+1}, \quad \text{if } s_n = 0, \tag{36}$$

while in case $s_n = 1$ we redefine $\tilde{G}_{n,n+1}$ by the same formula (36) but now with the factors multiplied in the reverse order. We can then put

$$\tilde{G}_{n-1,n} := I_2. \tag{37}$$

Note that after the above updates, the value of the snake-shaped matrix factorization $\mathcal{S}$ remains unchanged but we have succeeded to transform the Givens transformation $\tilde{G}_{n-1,n}$ in (34) into the identity matrix I. Then it is easily seen that the snake shape of $\mathcal{S}$ can be 'broken' into two pieces, in the sense that $\mathcal{S} = UV$ where $U = \prod_{k=0}^{n-2} G_{k,k+1}$ is the submatrix formed by rows and columns $0, \ldots, n-1$ of $\mathcal{S}$, while $V = \prod_{k=n}^{\infty} G_{k,k+1}$ is the submatrix formed by rows and columns $n, \ldots, \infty$. Note that the matrices U and V have a complementary zero pattern and hence they commute with each other.

Let us now denote with $\mathcal{S}_{n-1} := \prod_{k=0}^{n-2} G_{k,k+1}$ the topmost part of the 'broken' snake $\mathcal{S}$. Note that $\mathcal{S}_{n-1}$ is a snake-shaped matrix factorization of size $n \times n$; in particular it is still *unitary*. Note also that this matrix depends on the parameter $\theta \in \mathbb{R}$ by means of (35).

Remark 10. The principal $n \times n$ submatrix of $\mathcal{S}$ can be obtained in the same way as above, but now replacing the role of $\mathrm{e}^{\mathrm{i}\theta} \in \mathbb{T}$ in (34) by the original matrix entry $\overline{\alpha_n}$. Note however that $\overline{\alpha_n}$ lies strictly inside the unit disk and hence the resulting $n \times n$ submatrix is *not* unitary anymore; cf. the motivation earlier in this section.

One has then the following result.

Theorem 11 (*Gragg [18]*). *Let $\theta \in \mathbb{R}$ be fixed. Using the above construction, the eigenvalues of $\mathcal{H}_{n-1}$ are distinct, belong to $\mathbb{T}$ and appear as nodes in an n-point Szegő quadrature formula for the measure μ. The corresponding quadrature weights are the first components of the normalized eigenvectors of $\mathcal{H}_{n-1}$.*

Proposition 12 (*Watkins [33]*). *Theorem 11 also holds for the matrix $\mathcal{C}_{n-1}$.*

Here with 'normalized' eigenvectors we mean that the eigenvectors should be scaled in such a way that they form an orthonormal system and that their first components are real positive numbers.

The characteristic polynomial of the above matrix $\mathcal{H}_{n-1}$ (or equivalently, $\mathcal{C}_{n-1}$) is known as a monic *para-orthogonal polynomial* of degree n [23]. Note that this polynomial depends on the free parameter θ, and hence there is in fact a one-parameter *family* of para-orthogonal polynomials (and so, a one-parameter family of n-point Szegő quadrature formulas for μ).

Now one could ask why there is such a similarity between $\mathcal{H}$ and $\mathcal{C}$ in the above results. This is explained by the following basic observation, which is essentially due to Ammar, Gragg and Reichel [1] for the case of $\mathcal{H}$ and $\mathcal{C}$.

Proposition 13 (*Based on Ammar, Gragg and Reichel [1]*). *Let $\theta \in \mathbb{R}$ be fixed. Then the eigenvalues and the first components of the normalized eigenvectors of $\mathcal{S}_{n-1}$ depend on the Schur parameters but not on the shape of the snake.*

Proof. Recall that the snake-shaped matrix factorization is given by $\mathcal{S}_{n-1} = \prod_{k=1}^{n-2} G_{k,k+1}$, for some order of the factors. But it is a general fact that the matrices AB and BA have the same eigenvalues; this follows from the similarity transformation

$$AB \mapsto A^{-1}(AB)A = BA. \tag{38}$$

By applying this idea recursively for the choice $A = \prod_{k=l}^{n-2} G_{k,k+1}$, for $l = n-2, \ldots, 1$ (only those indices l for which $s_l = 1$ have to be treated), one can succeed to rearrange the Givens transformations of S_{n-1} into the unitary Hessenberg form $G_{0,1}G_{1,2}\cdots G_{n-2,n-1}$ (compare with (6)). It follows that the eigenvalues of $\mathcal{S}_{n-1}$ are indeed independent of the order of the factors $G_{k,k+1}$, i.e., they are independent of the shape of the snake.

The same argument also shows that the first components of the normalized eigenvectors are independent of the shape of the snake. To see this, consider the eigen-decomposition $\mathcal{S}_{n-1} = UDU^*$, where D is a diagonal matrix containing the eigenvalues, and U is a unitary matrix whose columns are the eigenvectors, scaled in such a way that the first row of U has real positive entries. The point is now that the only Givens transformation of $\mathcal{S}_{n-1}$ acting on the 0th index is $G_{0,1}$; but in the above argument the latter can only appear as the B-factor in (38), and hence the first row of U is easily seen to remain unchanged under the similarity (38). □

Corollary 14. *Theorems 8 and 11 hold with $\mathcal{H}$ replaced by any snake-shaped matrix factorization $\mathcal{S}$.*

Proof. This follows from Proposition 13 and Remark 10. □

Note that Proposition 13 implies that the eigenvalue problems for the matrices $\mathcal{H}_{n-1}, \mathcal{C}_{n-1}$ and $\mathcal{S}_{n-1}$ are conceptually equivalent. Interestingly, these problems turn out to be also *numerically* equivalent since, for reasons of efficiency and numerical stability, the eigenvalue computation for $\{\mathcal{H}_{n-1}, \mathcal{C}_{n-1}, \mathcal{S}_{n-1}\}$ should preferably be performed using their factorization as a product of Givens transformations, rather than using their entry-wise expansions.

Finally, we mention that the development of extensions of Szegő quadrature formulas and the investigation of the connection between them and Gauss quadrature formulas on the interval $[-1, 1]$ are active areas of research: see e.g. [3,11,12,22] and the references therein. A whole variety of practical eigenvalue computation algorithms for unitary Hessenberg and five-diagonal matrices has already been developed in the literature. In [26], Rutishauser designed an LR-iteration. Implicit QR-algorithms for unitary Hessenberg matrices were described and analyzed in [9,14,19,29]. In [2,21] and the references therein, divide and conquer algorithms were constructed. Other approaches are an algorithm using two half-size singular value decompositions [1], a method involving matrix pencils [4], and a unitary equivalent of the Sturm sequence method [5].

Acknowledgments

The authors thank professors A. Bultheel and A. Kuijlaars for their useful suggestions and the referees for their valuable comments.

The work of first author is partially supported by the research project MTM 2005-08571 of the Spanish Government and by the Fund of Scientific Research (FWO), project “RAM: Rational modeling: optimal conditioning and stable algorithms”, grant ♯G.0423.05 and the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with the author. The work of second author is supported by the Onderzoeksfonds K.U.Leuven/Research Fund K.U.Leuven.

References

[1] G.S. Ammar, W.B. Gragg, L. Reichel, On the eigenproblem for orthogonal matrices, in: 25th IEEE Conference on Decision and Control, Athens, Greece, 1963–1966, 1986.

[2] G.S. Ammar, L. Reichel, D.C. Sorensen, An implementation of a divide and conquer algorithm for the unitary eigenproblem, ACM Trans. Math. Software 18 (3) (1992) 292–307.

[3] A. Bultheel, L. Daruis, P. González-Vera, A connection between quadrature formulas on the unit circle and the interval $[-1, 1]$, J. Comput. Appl. Math. 132 (2001) 1–14.

[4] A. Bunste-Gerstner, L. Elsner, Schur parameter pencils for the solution of the unitary eigenproblem, Linear Algebra Appl. 154-156 (1992) 741–778.

[5] A. Bunse-Gerstner, C. He, On a Sturm sequence of polynomials for unitary Hessenberg matrices, SIAM J. Matrix Anal. Appl. 16 (4) (1995) 1043–1055.

[6] M.J. Cantero, R. Cruz-Barroso, P. González-Vera, A matrix approach to the computation of quadrature formulas on the unit circle, Appl. Numer. Math. 58 (3) (2008) 296–318.
[7] M.J. Cantero, L. Moral, L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, Linear Algebra Appl. 362 (2003) 29–56.
[8] M.J. Cantero, L. Moral, L. Velázquez, Minimal representations of unitary operators and orthogonal polynomials on the unit circle, Linear Algebra Appl. 408 (2005) 40–65.
[9] S. Chandrasekaran, M. Gu, J. Xia, J. Zhu, A fast QR algorithm for companion matrices, Recent Adv. Matrix Operator Theory 179 (2007) 111–143.
[10] R. Cruz-Barroso, P. González-Vera, A Christoffel-Darboux formula and a Favard's theorem for orthogonal Laurent polynomials on the unit circle, J. Comput. Appl. Math. 179 (2005) 157–173.
[11] R. Cruz-Barroso, P. González-Vera, Orthogonal Laurent polynomials and quadratures on the unit circle and the real half-line, Electron. Trans. Numer. Anal. 19 (2005) 113–134.
[12] R. Cruz-Barroso, L. Daruis, P. González-Vera, O. Njåstad, Sequences of orthogonal Laurent polynomials, bi-orthogonality and quadrature formulas on the unit circle, J. Comput. Appl. Math. 200 (2007) 424–440.
[13] S. Delvaux, M. Van Barel, Unitary rank structured matrices, J. Comput. Appl. Math. 215 (1) (2008) 49–78.
[14] S. Delvaux, M. Van Barel, Eigenvalue computation for unitary rank structured matrices, J. Comput. Appl. Math. 213 (1) (2008) 268–287.
[15] P.E. Gill, G.H. Golub, W. Murray, M.A. Saunders, Methods for modifying matrix factorizations, Math. Comp. 28 (1974) 505–535.
[16] L. Golinskii, M. Kudryavtsev, An inverse spectral theory for finite CMV matrices, (in press) arXiv:0705.4353.
[17] G.H. Golub, C.F. Van Loan, Matrix Computations, third edition, The Johns Hopkins University Press, 1996.
[18] W.B. Gragg, Positive definite Toeplitz matrices, the Arnoldi process for isometric operators and Gaussian quadrature on the unit circle, J. Comput. Appl. Math. 46 (1993) 183–198. This is a slightly revised version of a paper by the same author and published in Russian in: E. S. Nicholaev editor, Numer. Meth. Lin. Alg., Moscow University Press, Moscow (1982), 16–32.
[19] W.B. Gragg, The QR algorithm for unitary Hessenberg matrices, J. Comput. Appl. Math. 16 (1986) 1–8.
[20] U. Grenander, G. Szegő, Toeplitz Forms and Their Applications, Chelsea, New York, NY, 1984.
[21] M. Gu, R. Guzzo, X.-B. Chi, X.-Q. Cao, A stable divide and conquer algorithm for the unitary eigenproblem, SIAM J. Matrix Anal. Appl. 25 (2003) 385–404.
[22] C. Jagels, L. Reichel, Szegő-Lobatto quadrature rules, J. Comput. Appl. Math. 200 (2007) 116–126.
[23] W.B. Jones, O. Njåstad, W.J. Thron, Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle, Bull. London Math. Soc. 21 (1989) 113–152.
[24] R. Killip, I. Nenciu, CMV: the unitary analogue of Jacobi matrices, Comm. Pure Appl. Math. 60 (2007) 1148–1188.
[25] I. Nenciu, CMV matrices in random matrix theory and integrable systems: A survey, J. Phys. A: Math. Gen. 39 (28) (2006) 8811–8822.
[26] H. Rutishauser, Bestimmung der eigenwerte orthogonaler matrizen, Numer. Math. 9 (1966) 104–108.
[27] B. Simon, Orthogonal Polynomials on the Unit Circle. Part 1: Classical Theory, in: Amer. Math. Soc. Coll. Publ., vol. 54, Amer. Math. Soc, Providence, R.I, 2005.
[28] B. Simon, CMV matrices: Five years after, J. Comput. Appl. Math. 208 (1) (2007) 120–154.
[29] M. Stewart, An error analysis of a unitary Hessenberg QR algorithm, SIAM J. Matrix Anal. Appl. 28 (1) (2006) 40–67.
[30] G. Szegő, Orthogonal polynomials, in: Amer. Math. Soc. Coll. Publ., vol. 23, Amer. Math. Soc, Providence, R.I, 1975.
[31] W.J. Thron, L-polynomials orthogonal on the unit circle, in: A. Cuyt (Ed.), Nonlinear Methods and Rational Approximation, Reidel Publishing Company, Dordrecht, 1988, pp. 271–278.
[32] L. Velázquez, Spectral methods for orthogonal rational functions, J. Funct. Anal. 254 (4) (2008) 954–986.
[33] D.S. Watkins, Some perspectives on the eigenvalue problem, SIAM Rev. 35 (3) (1993) 430–471.